

Weekly Report

Lu Junhua

2016 年 11 月 28 日

Done

- Went to netease and discussed with the 运维, 外挂组. Introduced our work, and learn some knowledge about their work and Liu Yong team's work.
- Examined and recruited 2 students of infovis course (undergraduate).
- Wrote and refined a short essay for vag's WeChat Subscribe Account.

To do

- Fetch the data and assign exploration tasks to undergraduate students.
- A more detailed knowledge on the game data and detection of bot, with students in yuquan and netease.

网易交谈摘要

由于ppt还没到手, 有些事宜仍然需要确认下.

- 首先, 明确他们现在主要攻克的外挂类型(非常规挂. 常规挂一般基于客户端方法, 用一个类似于杀毒软件的就能解决大多数问题. 对经济收益以及游戏经济平衡影响不大, 不太关注. 非常规的就得用服务器方法).
- 此类型外挂利润相当可观, 运作也异常严谨. 一般完成任务现在是自动主线, 自动三环两种. 等级也有一定的特殊性(由任务决定, 而且由于是必经的一个过程). 耗时其实可能和正常人差不太多, 简单方法不容易判断. 游戏中使用外挂的比例并不小.
- 公司现在也采用基与行为的检测方法, 特征主要靠人来识别. 这是方法中比较不好的一个环节, 因为有时候要讲究时效性或者及时性. 游戏如果有改, 特征也变; 封了外挂, 外挂团队也会变.
- 调研发现他们并不会聊天, 也就是说不会有社交网络.
- 有一些考虑, 比如说我们能不能做一个游戏特征选择的工具, 只做分类之前的步骤; 如果神经网络对于这种隐藏特征能有很好的把握, 能用神经网络去做预测吗? 此外张量分解也是有类似作用的, 还能补足一些缺失数据(这种情况还是有的); 时间序列方面, 上次说的自相似性是否可以使用. 对于特征的及时性, 是不是考虑一些流数据或者并行处理之类的东西.

论文

- SIAM review 2009 *Tensor Decompositions and Applications* 关于张量分解综述, 从基本概念开始, 一些诸如fiber, Matricization(矩阵化), n-mode (matrix) 乘积, rank(秩)等有所了解.
- TVCG 2015 *1.5d Egocentric Dynamic Network Visualization* 老文章, 关于egonetwrok 的. 将时间作为一个单独维度, 并提供其他可选方案. 外面的点, 有向线段可以自定义layout, 赋予不同的含义. ego除了表示人, 还可表示事件的聚合. 这种方法虽然比较简陋, 但也是相对最直观的形式.
- VAST 2016 *DropoutSeer: Visualizing Learning Patterns in Massive Open Online Courses for Dropout Reasoning and Prediction* Mooc数据, 数据量不大, 先分类, 后聚类. 看每一聚类的特征, 可以动态调节参数等. 界面比较美观, 但也与去年的egoSlider比较相似. 看起来分类方法效果一般, case里体现不出太多预测的功能.
- TVCG/VAST 2016 *SemanticTraj: A New Approach to Interacting with Massive Taxi Trajectories* 轨迹语义化以后给予查询, 通过不同的方式(旅程文档轨迹文档)进行索引. 这个系统优点在于, 简介明了, 功能不多不少(相信有些功能平时我们也非常想用), 可视化编码合理不冗余不牵强. 没有高深的公式, 单纯做工具, 这个值得我们去思考.
- TVCG/VAST 2016 *Towards Better Analysis of Deep Convolutional Neural Networks* 本文也相当于一款工具, 帮助人在有一定相关知识基础上, 了解卷积神经网络. 本身神经网络很复杂, 怎么打开黑盒, 不破坏本身结构还要减少视觉混乱, 这篇文章就初步解决这些问题. 文章交互比较简单, 但也差不太多. 后面的case很好的用上了前面的设计, 展示系统的有效性.
- 另外学习了西瓜书上有关特征选择的内容, 之前虽有了解却没深入. 本书写明白过滤式、包裹式、嵌入式三种. 后面会结合一些之前可视化的作品, 自习了解现有作品在这三个方面上做的内容.